

[illegible]

Information Integration – Task 3

[illegible]

Kurzer Rückblick

- Uns stehen folgende Datenquellen zur Verfügung:
 - IHME: Allgemeine Daten zu Krankheiten in verschiedenen Ländern
 - GHDx: Daten zum Konsum von Tabakwaren
 - WDI: Allgemeine Bevölkerungsdaten (Einkommensschichten, ...)
- Ziel ist es, Korrelationen zwischen Einflussfaktoren auf den Konsum von Tabakwaren und der Sterblichkeitsrate aufzudecken



„Höhere Einschränkung von Werbungen zu Tabakwaren“ → „geringere Anzahl an Rauchern“

„Höhere Unterstützung, mit dem Rauchen aufzuhören“ → „geringere Anzahl an weibl. Rauchern“

„Höhere Steuern“ → „Wesentlich weniger Jugendliche, die rauchen“

→ jeweils Betrachtung der Auswirkungen auf die allgemeine Sterblichkeitsrate

→ **was sind die effektivsten Mittel, um die Sterblichkeitsrate zu verringern?**

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 2

Schema zu Beginn

WDI

country_id	Country Code
country_name	Country Name
indicator_name	Indicator Name
indicator_code	Indicator Code
unit	Unit
source	Source
year	Year
value	Value

IHME

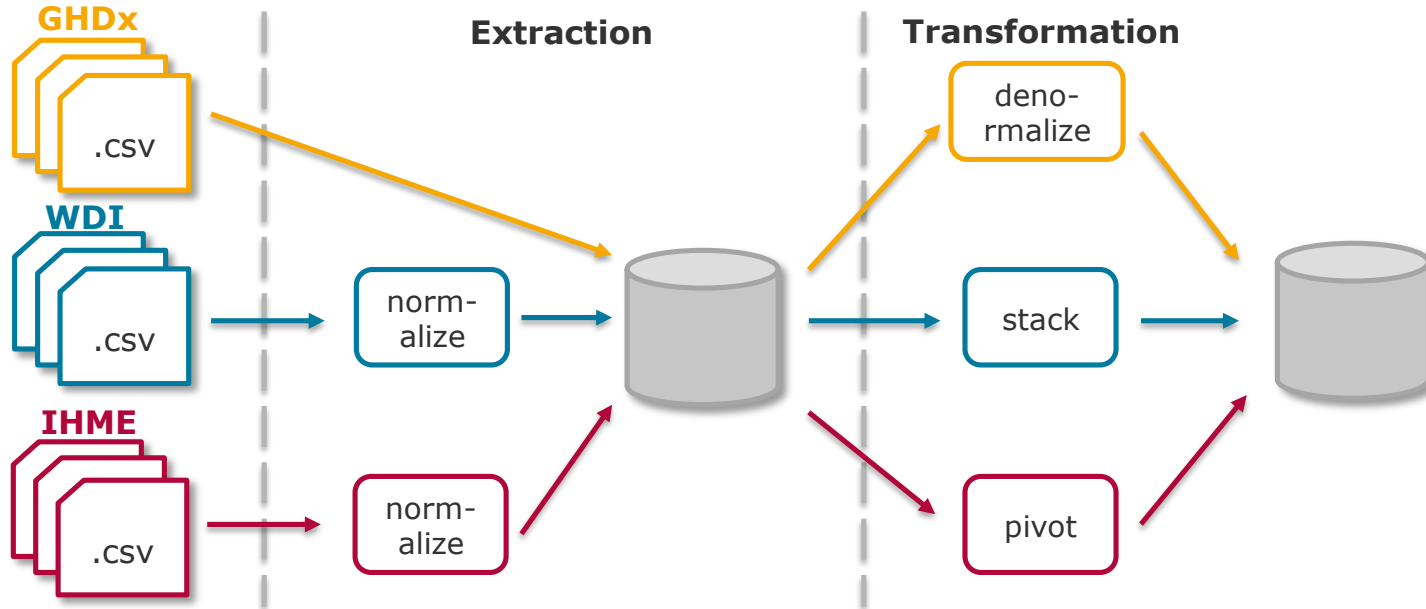
GHDx

country_id	Country Code
country_name	Country Name
indicator_name	Indicator Name
indicator_code	Indicator Code
unit	Unit
source	Source
year	Year
value	Value

Task 3 - Integration
 Sascha Obst,
 Johannes Hötter
 05.12.19
 Chart 3

Datenfluss in unserem System

Integration in unserer Datenbank ist theoretisch größtenteils über Länder-/Jahresschlüssel bereits gegeben; zur praktischen Nutzung müssen jedoch verschiedene Tabellen in ihrer Struktur beeinflusst werden:



Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 4

Finales globales Schema

WDI		
Country Code	text	
Short Name	text	
Table Name	text	
Long Name	text	
2-alpha code	text	
Currency Unit	text	
Special Notes	text	
Region	text	
Income Group	text	
WB-2 code	text	
National accounts base year	text	
National accounts reference year	text	
SNA price valuation	text	
Lending category	text	
Other groups	text	
System of National Accounts	text	
Alternative conversion factor	text	
PPP survey year	text	
Balance of Payments Manual in use	text	
External debt Reporting status	text	
System of trade	text	
Government Accounting concept	text	
IMF data dissemination standard	text	
Latest population census	text	
Latest household survey	text	
Source of most recent Income and expenditure data	text	
Vital registration complete	text	
Latest agricultural census	text	
Latest industrial data	double precision	
Latest trade data	double precision	
Unnamed: 30	double precision	

WDI		
CountryCode	text	
SeriesCode	text	
Year	text	
Unnamed: 4	double precision	
DESCRIPTION_ID	bigint	
key	bigint	
value	text	
CountryCode	text	
SeriesCode	text	
DESCRIPTION	text	
Unnamed: 3	double precision	

GHDx		
Entity	text	
Code	text	
Year	bigint	
70+ years old (deaths)	double precision	
15-49 years old (deaths)	double precision	
50-69 years old (deaths)	double precision	
Entity	text	
Code	text	
Year	bigint	
Adults who smoke (%)	double precision	
GDP per capita (international-\$) (constant 2011 international \$)	double precision	
Unnamed: 5	double precision	

IHME		
measure_id	bigint	
location_id	bigint	
sex_id	bigint	
age_id	bigint	
cause_id	bigint	
metric_id	bigint	
year	bigint	
val	double precision	
upper	double precision	
lower	double precision	
index	bigint	
key	bigint	
value	text	
index	bigint	
key	bigint	
value	text	
index	bigint	
key	bigint	
value	text	

GHDx		
Code	text	
Year	text	
Indicator:Average - cigarette price in international dollars (double precision	
Share of women (% of women)	double precision	
Share of men (% of men)	double precision	
Unnamed: 5	double precision	
Cigarette consumption per smoker per day (cigarettes)	double precision	
Estimated daily consumption (cigarettes)	double precision	
Upper bound (cigarettes)	double precision	
Lower bound (cigarettes)	double precision	
Upper bound (%)	double precision	
Lower bound (%)	double precision	
Estimated prevalence (%)	double precision	
Indicator:Enforce bans on tobacco advertising	double precision	
Secondhand smoke	double precision	
Tobacco smoking	double precision	
Number of daily smokers - both (IHME, GHDx (2012))	double precision	
Sales of cigarettes per adult per day (International Smoking St	double precision	
70+ years old (deaths)	double precision	
5-14 years old (deaths)	double precision	
15-49 years old (deaths)	double precision	
50-69 years old (deaths)	double precision	
Under-5s (deaths)	double precision	
Smoking prevalence, total (ages 15+) (% of adults)	double precision	
Age-standardized share of cancer deaths attributed to tobacco	double precision	
Indicator:Raise taxes on tobacco	double precision	
Smoking (deaths)	double precision	
Secondhand smoke (deaths)	double precision	
Indicator:Offer help to quit tobacco use	double precision	
Indicator:Average - taxes as a % of cigarette price - total tax	double precision	

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 5

Was wir gelernt haben

- Vermeintlich schwierige Aufgaben können sehr einfach sein ...
 - IHME-Daten: 31 Tabellen mit selbem Schema
 - Nach Recherche sind von den 16.500.000 Datensätzen noch 950.000 relevant
 - Über UNION verknüpft und gespeichert, Rest gelöscht
- ... und vermeintlich einfache Aufgaben können sehr schwer werden!
 - GHDx-Daten: 17 Tabellen mit selbem Schlüssel und jeweils einer Kennzahl („Code“ + „Year“ und z.B. „Share Women“)
 - Aber: kaum Schlüssel, die in allen 17 Tabellen vorkommen -> INNER JOIN nicht möglich, also FULL OUTER JOIN
 - Aber: Schlüssel geht verloren, wenn aus einer festen Spalte der 17 Tabellen bezogen, in welcher der Schlüssel nicht vorhanden ist
 - Also: je Schlüssel eine Anfrage mit FULL OUTER JOIN und Schlüssel als „dynamisch gesetzte Konstante“ ausführen -> extrem aufwändig!

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 6

■ WDI-Daten: Pivottisierung notwendig!

	"C..."	"Indicato..."	"1960"	"1961"	"1962"	"1963"	"1964"	"1965"	"1966"	"1967"
1	Arab World	ARB	2005 PPP conversi...	PA.NUS...	<null>	<null>	<null>	<null>	<null>	<null>
2	Arab World	ARB	2005 PPP conversi...	PA.NUS...	<null>	<null>	<null>	<null>	<null>	<null>
3	Arab World	ARB	Access to clean f...	EG.CFT...	<null>	<null>	<null>	<null>	<null>	<null>
4	Arab World	ARB	Access to electri...	EG.ELC...	<null>	<null>	<null>	<null>	<null>	<null>
5	Arab World	ARB	Access to electri...	EG.ELC...	<null>	<null>	<null>	<null>	<null>	<null>
6	Arab World	ARB	Access to electri...	EG.ELC...	<null>	<null>	<null>	<null>	<null>	<null>
7	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
8	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
9	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
10	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
11	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
12	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
13	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
14	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
15	Arab World	ARB	Account ownership...	FX.OWN...	<null>	<null>	<null>	<null>	<null>	<null>
16	Arab World	ARB	Adequacy of socia...	per_si...	<null>	<null>	<null>	<null>	<null>	<null>
17	Arab World	ARB	Adequacy of socia...	per_all...	<null>	<null>	<null>	<null>	<null>	<null>
18	Arab World	ARB	Adequacy of socia...	per_sa...	<null>	<null>	<null>	<null>	<null>	<null>
19	Arab World	ARB	Adequacy of unemp...	per_lm...	<null>	<null>	<null>	<null>	<null>	<null>
20	Arab World	ARB	Adjusted net enro...	SE.PR...	<null>	<null>	<null>	<null>	<null>	<null>
21	Arab World	ARB	Adjusted net enro...	SE.PR...	<null>	<null>	<null>	<null>	<null>	<null>
22	Arab World	ARB	Adjusted net enro...	SE.PR...	<null>	<null>	<null>	<null>	<null>	<null>
23	Arab World	ARB	Adjusted net nati...	NY.ADJ...	<null>	<null>	<null>	<null>	<null>	<null>
24	Arab World	ARB	Adjusted net nati...	NY.ADJ...	<null>	<null>	<null>	<null>	<null>	<null>
25	Arab World	ARB	Adjusted net nati...	NY.ADJ...	<null>	<null>	<null>	<null>	<null>	<null>

	"Country Code"	"Indicator Code"	variable	value
1	ARB	SP.ADO.TFRT	1960	134.772340547077
2	ARB	SP.POP.DPND	1960	88.0611098786549
3	ARB	SP.POP.DPND.OL	1960	6.59195099288584
4	ARB	SP.POP.DPND.YG	1960	81.3241863909366
5	ARB	ER.FSH.AQUA.MT	1960	4600
6	ARB	MS.MIL.XPRT.KD	1960	30000000
7	ARB	MS.MIL.MPRT.KD	1960	539000000
8	ARB	SP.DYN.CBRT.IN	1960	47.7900762443202
9	ARB	FM.LBL.BMNY.GD.ZS	1960	22.0737099987733
10	ARB	ER.FSH.CAPT.MT	1960	474880
11	ARB	FS.AST.CGVOV.GD.ZS	1960	0.401705949228533


```
df = pd.read_sql_table('wdidata', con=engine)
df = pd.melt(df, id_vars=['Country Code', 'Indicator Code'],
            value_vars=[str(val) for val in range(1960, 2019, 1)]).dropna()
df_to_sql(df, 'wdidata_pivot', engine, if_exists='replace')
```

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 7

Beispiele

- IHME-Daten: Einschränkung der relevanten Wertebereiche, dadurch Reduzierung der Tabellen

```
# filter values for the ihme tables
# only rows containing information about the following topics will be addressed
# reduces ihme content from 360 diseases to the 27 relevant for smoking
filter_values = [
    "= 'Esophageal cancer'",
    "= 'Liver cancer'",
    "= 'Ectopic pregnancy'",
    "= 'Tuberculosis'",
    "= 'Rheumatoid arthritis'",
    "= 'Bladder cancer'",
    "= 'Cervical cancer'",
    "= 'Psoriasis'",
    "LIKE '%diabetes%'",
    "LIKE '%heart%'",
    "LIKE '%vision%'",
    "LIKE '%lung%'",
    "LIKE '%stroke%'",
    "LIKE '%hearing%'",
    "LIKE '%dementia%'"
]
```

Werte recherchiert (Arztbefragung, Artikel)

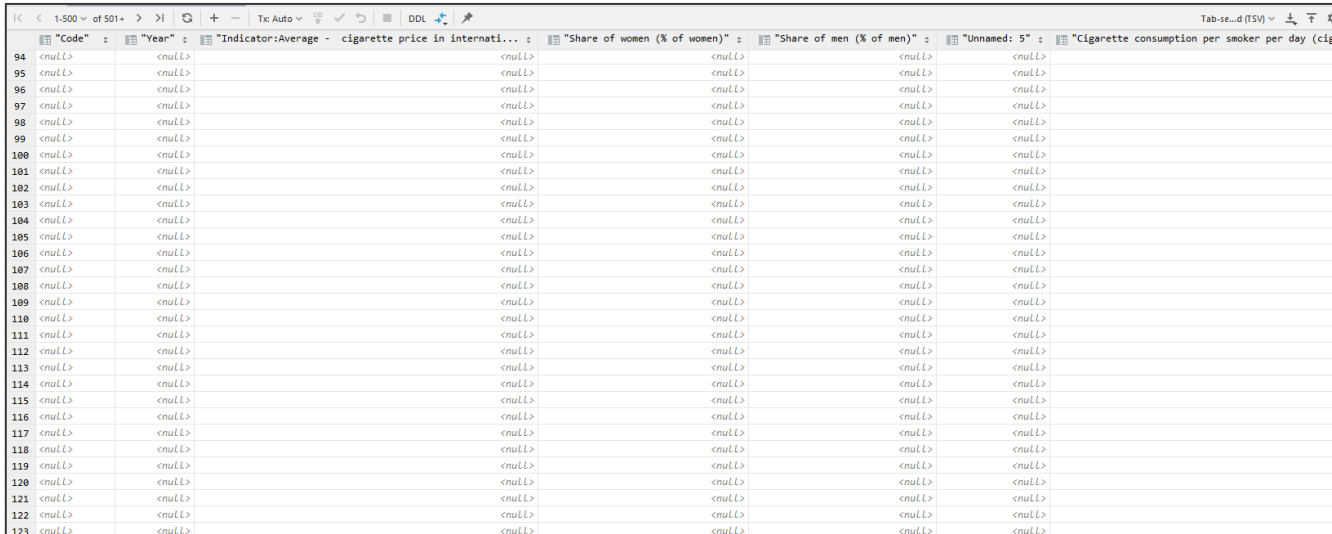
	index	key	value
1	0	504	Non-rheumatic valvular heart disease
2	1	487	Leukemia
3	2	494	Stroke
4	3	484	Hodgkin lymphoma
5	4	485	Non-Hodgkin lymphoma
6	5	486	Multiple myeloma
7	6	489	Other malignant neoplasms
8	7	490	Other neoplasms
9	8	491	Cardiovascular diseases
10	9	492	Rheumatic heart disease
11	10	493	Ischemic heart disease
12	11	495	Ischemic stroke
13	12	496	Intracerebral hemorrhage
14	13	497	Subarachnoid hemorrhage
15	14	970	Other non-rheumatic valve diseases
16	15	498	Hypertensive heart disease
17	16	968	Non-rheumatic calcific aortic valve disease
18	17	969	Non-rheumatic degenerative mitral valve disease
19	18	971	Cirrhosis due to NASH
20	19	973	Substance use disorders
21	20	974	Diabetes and kidney diseases
22	21	975	Diabetes mellitus type 1
23	22	976	Diabetes mellitus type 2
24	23	980	Bacterial skin diseases
25	24	992	Upper digestive system diseases
26	25	995	Maternal abortion and miscarriage

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 8

Beispiele

- GHDx-Daten: „Das sind zwar 17 Tabellen, aber die verknüpfen wir dann einfach miteinander über deren Key, das passt schon...“
 - So einfach war es dann doch nicht!



Code	Year	Indicator: Average - cigarette price in internati...	Share of women (% of women)	Share of men (% of men)	Unnamed; 5"	Cigarette consumption per smoker per day (cigi
94	<null>	<null>	<null>	<null>	<null>	<null>
95	<null>	<null>	<null>	<null>	<null>	<null>
96	<null>	<null>	<null>	<null>	<null>	<null>
97	<null>	<null>	<null>	<null>	<null>	<null>
98	<null>	<null>	<null>	<null>	<null>	<null>
99	<null>	<null>	<null>	<null>	<null>	<null>
100	<null>	<null>	<null>	<null>	<null>	<null>
101	<null>	<null>	<null>	<null>	<null>	<null>
102	<null>	<null>	<null>	<null>	<null>	<null>
103	<null>	<null>	<null>	<null>	<null>	<null>
104	<null>	<null>	<null>	<null>	<null>	<null>
105	<null>	<null>	<null>	<null>	<null>	<null>
106	<null>	<null>	<null>	<null>	<null>	<null>
107	<null>	<null>	<null>	<null>	<null>	<null>
108	<null>	<null>	<null>	<null>	<null>	<null>
109	<null>	<null>	<null>	<null>	<null>	<null>
110	<null>	<null>	<null>	<null>	<null>	<null>
111	<null>	<null>	<null>	<null>	<null>	<null>
112	<null>	<null>	<null>	<null>	<null>	<null>
113	<null>	<null>	<null>	<null>	<null>	<null>
114	<null>	<null>	<null>	<null>	<null>	<null>
115	<null>	<null>	<null>	<null>	<null>	<null>
116	<null>	<null>	<null>	<null>	<null>	<null>
117	<null>	<null>	<null>	<null>	<null>	<null>
118	<null>	<null>	<null>	<null>	<null>	<null>
119	<null>	<null>	<null>	<null>	<null>	<null>
120	<null>	<null>	<null>	<null>	<null>	<null>
121	<null>	<null>	<null>	<null>	<null>	<null>
122	<null>	<null>	<null>	<null>	<null>	<null>
123	<null>	<null>	<null>	<null>	<null>	<null>

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 9

- Nach mehreren Anläufen entstand folgendes Skript:

```
select_map = defaultdict(list)
for table in table_names:
    header_df = pd.read_sql_query("SELECT * FROM {} LIMIT 1".format(table), con=engine).get_first_row_to_reflect
    select_map[table].extend(header_df.columns[3:])

alias_map = {}
alphabet = list('abcdefghijklmnopqrstuvwxyz')
for table in select_map.keys():
    alias_map[table] = alphabet.pop()

sql_select = ''
sql_join = ''
sql_condition = ''
last_alias = None # not last alias that exist, but last alias that has been used

# melting the rows together
for code, year in code_year_combinations_df.values:
    sql = text('SELECT \'{code}\' "Code",\n\t \'{year}\' "Year",'.format(code=code, year=year) +
              sql_template.replace('<CODE>', str(code)).replace('<YEAR>', str(year)))
    df = pd.read_sql_query(sql, con=engine)
    row = {}
    for col in df.columns:
        try:
            row[col] = np.max(df[col]) # there is only one row with a value, so max can be applied
        except:
            row[col] = np.nan
    df = pd.DataFrame(data=[row], columns=df.columns)
    df_to_sql(df, 'ghdx_measures', engine, if_exists='append')

# build the select part
sql_select += ', '.join(['\n\t{}'.format(alias, column) for column in columns]) + ','

# build the from/join part
if last_alias: # OUTER JOIN
    sql_join += '\nFULL OUTER JOIN {table} {alias}\n\tON ' \
               '{alias}."Code" = {last_alias}."Code" ' \
               'AND {alias}."Year" = {last_alias}."Year"'.format(table=table,
                                                                alias=alias,
                                                                last_alias=last_alias)
    sql_condition += '\n\tOR ({alias}."Code" = \'{CODE}>' AND {alias}."Year" = \'{YEAR}>')'.format(alias=alias, last_alias=last_alias)
else: # FROM
    sql_join = "\nFROM {}".format(table, alias)
    sql_condition += '\n\tWHERE ({alias}."Code" = \'{CODE}>' AND {alias}."Year" = \'{YEAR}>')'.format(alias=alias, last_alias=last_alias)
    last_alias = alias

sql_template = sql_select[:-1] + sql_join + sql_condition #[:-1] to get rid of last comma
```

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 10

■ Folgende Abfrage pro Land-/Jahreskombination einmal ausgeführt!

```
SELECT 'TKM' "Code",
       '1838' "Year",
       z."Indicator:Average - cigarette price in international dollars ("
       y."Share of women (% of women)",
       y."Share of men (% of men)",
       y."Unnamed: 5",
       x."Cigarette consumption per smoker per day (cigarettes)",
       w."Estimated daily consumption (cigarettes)",
       w."Upper bound (cigarettes)",
       w."Lower bound (cigarettes)",
       v."Upper bound (%)",
       v."Lower bound (%)",
       v."Estimated prevalence (%)",
       u."Indicator:Enforce bans on tobacco advertising",
       t."Secondhand smoke",
       s."Tobacco smoking",
       r."Number of daily smokers - both (IHME, GHDx (2012))",
       p."Sales of cigarettes per adult per day (International Smoking St",
       o."70+ years old (deaths)",
       o."5-14 years old (deaths)",
       o."15-49 years old (deaths)",
       o."50-69 years old (deaths)",
       o."Under-5s (deaths)",
       n."Smoking prevalence, total (ages 15+) (% of adults)",
       m."Age-standardized share of cancer deaths attributed to tobacco ("
       l."Indicator:Raise taxes on tobacco",
       k."Smoking (deaths)",
       k."Secondhand smoke (deaths)",
       j."Indicator:Offer help to quit tobacco use",
       i."Indicator:Average - taxes as a % of cigarette price - total tax"
FROM average_price_of_a_pack_of_cigarettes z
FULL OUTER JOIN comparing_the_share_of_men_and_women_who_are_smoking y
  ON y."Code" = z."Code" AND y."Year" = z."Year"
FULL OUTER JOIN consumption_per_smoker_per_day x
  ON x."Code" = y."Code" AND x."Year" = y."Year"
FULL OUTER JOIN consumption_per_smoker_per_day_bounds w
  ON w."Code" = x."Code" AND w."Year" = x."Year"
FULL OUTER JOIN daily_smoking_prevalence_bounds v
  ON v."Code" = w."Code" AND v."Year" = w."Year"
FULL OUTER JOIN enforcement_of_bans_on_tobacco_advertising u
  ON u."Code" = v."Code" AND u."Year" = v."Year"
FULL OUTER JOIN number_of_deaths_from_secondhand_smoke t
  ON t."Code" = u."Code" AND t."Year" = u."Year"
FULL OUTER JOIN number_of_deaths_from_tobacco_smoking s
  ON s."Code" = t."Code" AND s."Year" = t."Year"
FULL OUTER JOIN number_of_total_daily_smokers r
  ON r."Code" = s."Code" AND r."Year" = s."Year"
FULL OUTER JOIN sales_of_cigarettes_per_adult_per_day p
  ON p."Code" = r."Code" AND p."Year" = r."Year"
FULL OUTER JOIN secondhand_smoke_deaths_by_age o
  ON o."Code" = p."Code" AND o."Year" = p."Year"
FULL OUTER JOIN share_of_adults_who_smoke n
  ON n."Code" = o."Code" AND n."Year" = o."Year"
FULL OUTER JOIN share_of_cancer_deaths_attributed_to_tobacco m
  ON m."Code" = n."Code" AND m."Year" = n."Year"
FULL OUTER JOIN share_of_tobacco_retail_price_that_is_tax l
  ON l."Code" = m."Code" AND l."Year" = m."Year"
FULL OUTER JOIN smoking_and_secondhand_deaths k
  ON k."Code" = l."Code" AND k."Year" = l."Year"
FULL OUTER JOIN support_to_help_to_quit_tobacco_use j
  ON j."Code" = k."Code" AND j."Year" = k."Year"
FULL OUTER JOIN taxes_as_share_of_cigarette_price i
  ON i."Code" = j."Code" AND i."Year" = j."Year"
WHERE (z."Code" = 'USA' AND z."Year" = '1980')
OR (y."Code" = 'USA' AND y."Year" = '1980')
OR (x."Code" = 'USA' AND x."Year" = '1980')
OR (w."Code" = 'USA' AND w."Year" = '1980')
OR (v."Code" = 'USA' AND v."Year" = '1980')
OR (u."Code" = 'USA' AND u."Year" = '1980')
OR (t."Code" = 'USA' AND t."Year" = '1980')
OR (s."Code" = 'USA' AND s."Year" = '1980')
OR (r."Code" = 'USA' AND r."Year" = '1980')
OR (p."Code" = 'USA' AND p."Year" = '1980')
OR (o."Code" = 'USA' AND o."Year" = '1980')
OR (n."Code" = 'USA' AND n."Year" = '1980')
OR (m."Code" = 'USA' AND m."Year" = '1980')
OR (l."Code" = 'USA' AND l."Year" = '1980')
OR (k."Code" = 'USA' AND k."Year" = '1980')
OR (j."Code" = 'USA' AND j."Year" = '1980')
OR (i."Code" = 'USA' AND i."Year" = '1980')
```

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 11

Beispiele

- Dann aber doch noch mit dem Happy End! 😊

	"Code"	"Year"	"Indicator:Average - cigarette price in internati..."	"Share of women (% of women)"	"Share of men (% of men)"
1	MNG	2012	0	6	48.6
2	ZWE	2012	0.67	1.9	31.3
3	KHM	2012	0.86	2.7	38.3
4	KHM	2014	0.92	2.3	35.7
5	PAK	2012	0.97	3.8	37.3
6	ZWE	2014	1.07	1.7	31
7	PAK	2014	1.1	3.3	37
8	PHL	2012	1.23	8.7	44.2
9	KAZ	2012	1.26	7.8	46
10	BGD	2012	1.38	1.5	47
11	KAZ	2014	1.45	7.4	44.7
12	RUS	2012	1.52	22.9	60.6
13	BGD	2014	1.54	1.2	45.5
14	ARE	2012	1.55	1.4	36.1
15	GEO	2012	1.57	5.5	57
16	SLE	2014	1.58	10.2	42.9
17	SEN	2012	1.61	0.6	17.6
18	UZB	2012	1.64	1.4	25.9
19	COL	2012	1.67	6	16
20	ZMB	2012	1.68	3.6	26.1
21	ARM	2012	1.69	1.7	55.1
22	KEN	2012	1.69	1.5	22.5

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 12

Was wir gelernt haben

- Manchmal hilft es, die Doku genauer zu studieren:
 - % werden in PostgreSQL für Wildcards benutzt (mehrere Zeichen)
 - % werden in SQLAlchemy (Datenbank-Connection für Python) für String-Formatierungen verwendet
 - Konflikt sorgt für verwirrende Errors („can't interpret dict“)
 - SQLAlchemy bietet eine Wrapper-Funktion `text()`, mit der solche Hürden umgangen werden können

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart **13**

Nächste Schritte

■ Bereinigung der Daten!

Share of women (% of women) :	Share of men (% of men) :	Unnamed: 5
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32800
<null>	<null>	32807
<null>	<null>	32823
<null>	<null>	32846
<null>	<null>	32878
<null>	<null>	32917
<null>	<null>	32965
<null>	<null>	33020
<null>	<null>	33083

System of National Accounts
Country uses the 1993 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
<null>
Country uses the 1993 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology
Country uses the 1993 System of National Accounts methodology
Country uses the 2008 System of National Accounts methodology

■ Views erstellen, in denen Spalten umbenannt werden

Task 3 - Integration

Sascha Obst,
Johannes Hötter
05.12.19
Chart 14



Danke für die Aufmerksamkeit!

Sascha Obst, Johannes Hötter

HBRSlers

05.12.19